A Human-Centered Design and Evaluation Framework for Information Search

Yang Gong, MD, MS, Jiajie Zhang, PhD
School of Health Information Sciences
The University of Texas Health Science Center at Houston

Information search in a distributed environment is an interactive process between the user and the artifact. How the information is distributed across the user and the artifact determines the efficacy of information search. Using a human-centered method, UFuRT, we developed an information search model and a taxonomy of search tasks. Further, we developed prototypes to investigate the relationship between search tasks and interface types. Preliminary results of analysis reveal the requirement of distributed information for search tasks and help understand the complexity of the tasks in different types of interfaces.

1. Introduction

Information plays a central role in organizational life and everyday activities. Searching information often involves many steps and requires tools [1-6]. How the information is distributed across a user's head (internal representations) and information tools (external representations) affects the efficiency of information search[7-9].

Distributed cognition plays a special role in understanding the interactions between people and technologies[10]. According to the theory of distributed cognition, cognitive activities are distributed across internal human minds, external cognitive artifacts, groups of people, and across space and time [11, 12]. Norman [12] argued that knowledge may be as much in the world as it is in the head. Therefore, the information carried by artifacts is as important to the achievement of a task as the knowledge residing in the mind of the artifact user.

Applying the theory of distributed cognition to information search which is a cognitive process, it may involve the coordination between internal and external (material or environmental) structures. When searching information in computers, the information displayed by the user interface and the information in the user's memory jointly determines the performance level of the search task.

According to Zhang's theory[7], external information presented in an appropriate format can reduce the difficulty of a task by supporting recognition-based memory or perceptual judgments rather than recall. Proper design of the user interface for information search can substantially increase the efficiency of human-computer interaction in terms of increased task performance, user satisfaction, user's knowledge retention, and decreased training time and error rate[13]. In many tasks, such as the search tasks in our study, people often use external artifacts to enhance internal memory and the artifacts are often created specifically for the purpose of remembering. For example, a patient chart is designed for reviewing a patient's medical history. Proper external representations support internal memories and therefore, enhance task performance.

In order to improve the quality of information search, it is important to understand how internal and external information interact with each other. During a distributed information search task, an information seeker often needs to match the information in his head with the information presented on the artifact. A typical task in an Electronic Health Record (EHR) system could be, for example, finding all the abnormal values of the lipid panels of a patient over the past 12 months. In this task, the normal range required, if not presented on screen, is the internal information. The observed values, which are presented in the patient's record, are the external information. How the information is distributed between internal and external representation determines the information search performance[14].

This study aims at developing an information search model and a taxonomy of search tasks under a human-centered theoretical framework which will help describe and explain the difference of the information search performance in different representations. We are also interested in the relationship between types of interfaces and types of search tasks in terms of effectiveness and efficiency for information search. We have two specific hypotheses. First, information search with more external information will yield a better task performance than those with less external information. Second, depending on the nature of tasks, the type of an interface affects search performance in different ways.

2. Information Models

Information search efficiency can be improved by several factors that characterize human information behaviors. These factors include choice of information sources, searching strategies, methods of verification of information reliability, and correspondence with earlier data[1]. Since distributed cognition provides an effective theoretical foundation for understanding human-computer interaction and is a useful framework for designing and evaluating information searching tools[11, 12], the focus of this research is on cognitive factors and their implications on human-computer interaction.

Studies on factors that affect human needs and information search behavior focused on the process of query formulation, execution, and results evaluation [3, 6, 15-18]. Existing information search models define a search task from the need of information seekers as well as the evaluation of results [6, 18, 19]. When these models are applied to a domain, such as healthcare, with information overload, time-pressure, and stress, they cannot adequately address the distributive, interactive nature of information search. In models designed for complex domains whereby the users and their tasks have a multifarious and rich nature, the distributed information resources should be considered [17]. The information search model proposed here is from the resource perspective. The choosing of proper information search resources is vital in time critical, complex information systems because proper resources or tools provide the required information in a timely manner. This also results from the interaction between internal and external information.

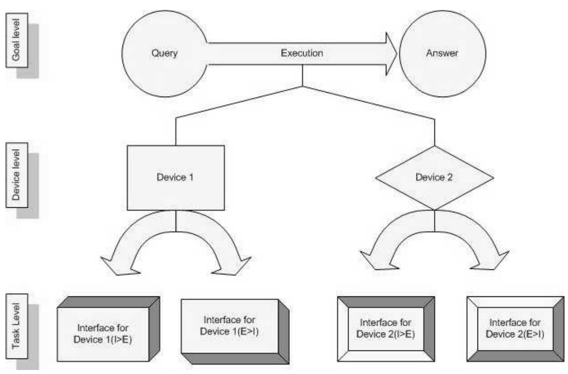

Figure 1. Three-level Information Searching Model

The proposal is of a three-level model (Figure 1) for information search. The three levels are goals, devices, and tasks. Different types of information search tasks may require different internal and/or external information depending on the nature of the device and the task. This model indicates that the source selection is dependent on the pattern of information distribution during the execution stage in an information search task. The model was not developed to replace any existing models of information search. Rather it adds the distributed information aspect to these models and fits well into them.

Existing information models can be categorized into four levels as follows (Figure 2):

Level one: Information behavior models. Wilson's model is a representative model in this category. It focuses on causes, consequences, and relationships among stages of an information-search activity [20]. Allen's gatekeeper model [21] refers to "a small number of key people to whom others frequently turned for information. These key people differed from their colleagues in the degree to which they exposed themselves to sources of technological information outside their organization." In brief, information behavior models focus on sources and channels of information.

Level two: Information seeking models are particularly concerned with the variety of methods people use to discover and gain access to information resources. Representative models at this level include Dervin's sense-making theory [22], Ellis's behavioral strategies [23, 24] and Kuhlthau's stages of information search model [25]. At this level, information seeking is considered as a purposive seeking to satisfy the need to achieve a goal.

Level three: Information searching models are particularly concerned with the interactions between user and computer-based information systems[26]. Our current study on distributed information falls into this category.

Level four: Visual search models are about the cognitive strategies that people use on specific displays. People perform visual search in parallel, sequential, and/or mixed search methods. The models at this level could be of help in explaining the information search performance in terms of the patterns of information distributions.

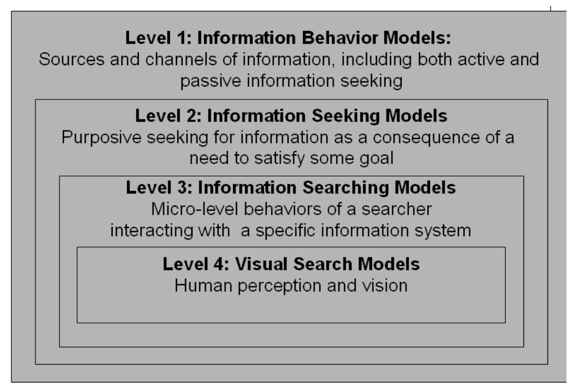

Figure 2. Existing information models

3. Methods

Zhang et al developed a method called *UFuRT* (User, Function, Representation, Task) for the design and evaluation of human-centered distributed information system[27]. *UFuRT* is built upon the theory of distributed cognition and a set of analysis techniques for human-computer system design. It emphasizes functions, users, tasks, and representations as indispensable components of a human-centered information system design. It provides systematic principles, guidelines, and procedures for designing human-centered information systems.

Theoretically, an information search interface designed by the *UFuRT* process ensures that the design matches the information search task which leads to a better task performance. Figure 3 shows the *UFuRT* procedures and their relations. In addition to the *UFuRT* process which establishes the mapping of users, functions, representations, and tasks in an information system, we also used GOMS (Goal-Operator-Method-Selection) task analysis to reveal information distributions for search tasks. GOMS analysis is a widely accepted method for analyzing human-computer interactions[28]. Combining GOMS with distributed cognition analysis provides us a unique perspective on the internal and external information required in each step of information search. The results from these analyses will be used to predict the efficacy of search tasks.

In brief, the *UFuRT* process enables us to uncover the search flows, steps, and internal and external information requirements for a search task. The process decomposes each search task into a finer granularity per interface whereby it can both explain and predict the search efficacy.

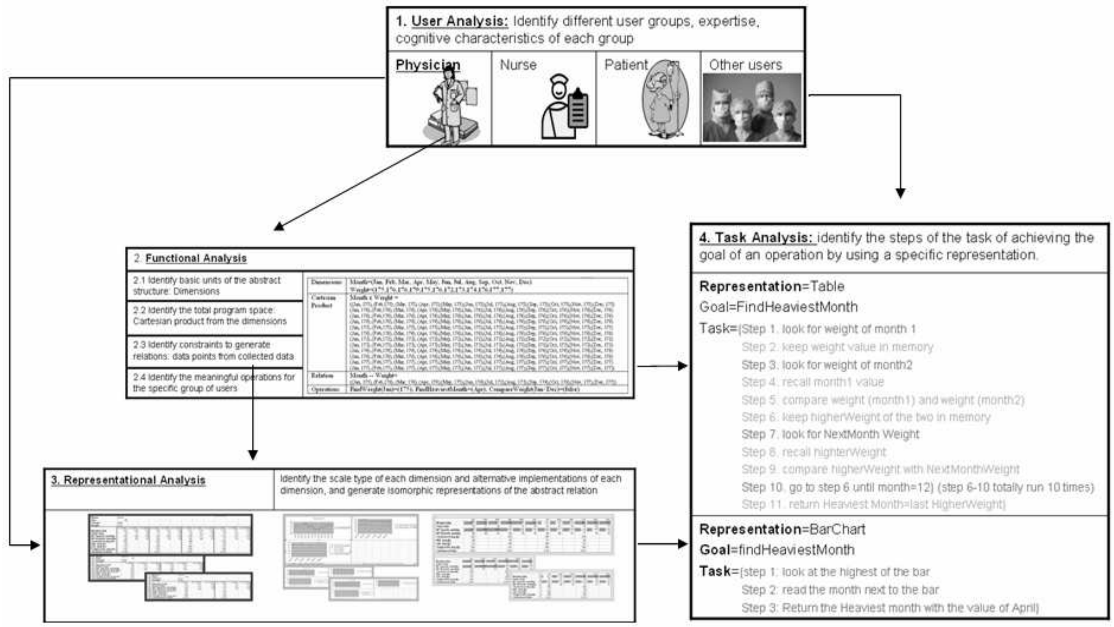

Figure 3. An Example of the *UFuRT* process (Based on Zhang et al.[27])

4. Development of Prototypes

Utilizing *UFuRT* to the distributed information search interface design, we developed several information search prototypes for the flowsheet module of an EHR system. According to the functional analysis, a record of a patient constructs many relations of the represented information. An information searching task is a process to search for specific information in the relations/dimensions. Search tasks can be categorized into direct search and comparative search. A direct search is to find a specific value under specific conditions. A direct search task could further be divided into a dimensional search or relational search. A comparative search compares the value within one dimension (within-dimension search) or between two (between-dimension search). Examples for each type of search are shown in Table 1.

Table 1. A taxonomy of information search tasks

	Direct Search		Comparative search	
	Dimensional Search	**Relational Search**	**Within-dimension**	**Between-dimension**
Definition	Search for values on one dimension	Search for values on multiple dimensions	Compare values within one dimension	Compare values between multiple dimensions
Example	Are there any abnormal levels of cholesterol in the patient record?	In which month of 2003, the patient's LDL level was abnormal?	Did the patient's triglyceride level drop since the start of his diet treatment?	Did the cholesterol ratio (total cholesterol/HDL) change over the past year?

The typical search task in this research requires both internal and external information. Applying data scale types to the taxonomy, each task has a set of operations which can be legitimately applied to data on the scale type. For example, a nominal search task in one dimensional search may be: Is there any abnormal cholesterol value in the patient's record? For answering this question, the internal information required is the normal cholesterol range (<200mg/dL) when it is not shown on the interface, while the observed values on the patient's chart are the external information. The interaction of the internal and external information may result in a nominal scale data, i.e. a "yes or no" answer. In a similar way, the other search tasks request the operations in ordinal, interval, or ratio scales.

In the experiments, tasks of dimensional search, relational search, and tasks within-dimension search will be conducted in three prototypes. These three prototypes we developed are based on some EHR product having both holistic views and separate views. The holistic view carries a 12-month health data while the separate view contains two 6-month health data on separate pages. In order to browse or compare the values between the first 6 months and second 6 months, a physician has to memorize the displayed values to compare. Therefore, the information search process requires the internal information be constrained by the design of the interface. Likewise, the graphic view and mixed view both have holistic and separate subtypes. Because of the different representations, the search method for the same task in a text display is dissimilar to the one in a graph representation. For example, to search for the heaviest weight in a year in the graph or mixed representation, one just needs to look at the highest bar (point) in the graph in order to get the corresponding month value. Consequently, this type of task in a graph display may have a better efficacy than text display (step 4 in Figure 3).

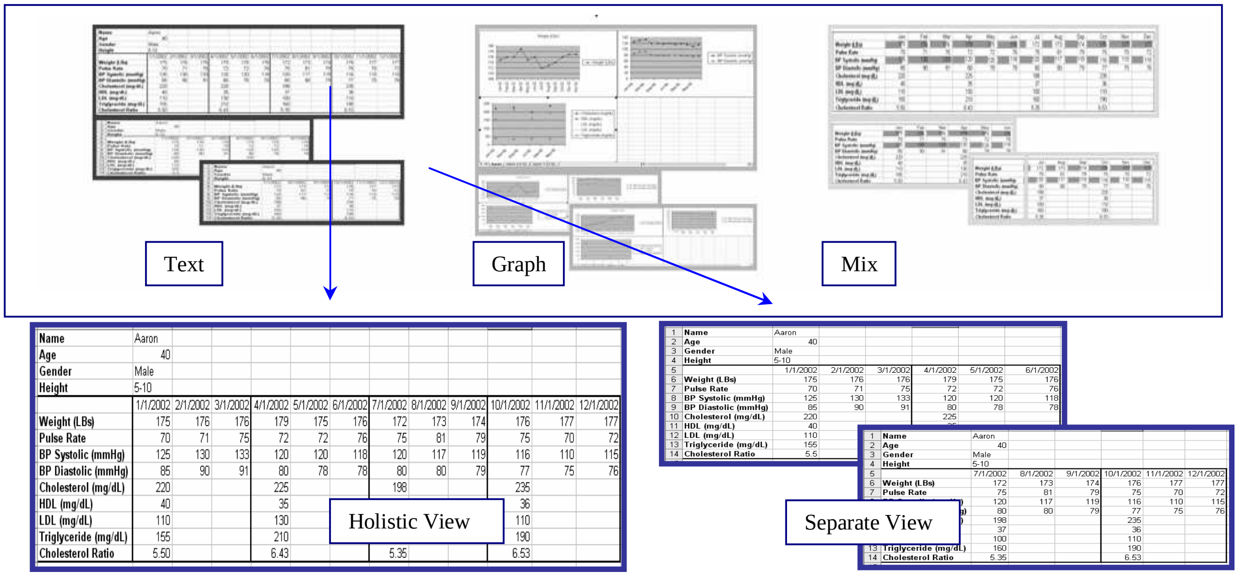

Figure 4. Three types of interfaces with holistic view and separate view

5. Preliminary Results

During the employment of the *UFuRT* theoretical framework to information search tasks per presentations, we have investigated three types of search tasks performed on three representations. Table 2 shows the task analysis results including task steps and internal/external information requirements. It not only explains the complexity of each search task but also describes the search performance on each interface. For instance, task II takes the most steps on average and task II's internal/external ratio is also the highest in the graph interface. Thus, according to this analysis, task II is the hardest search task among the three.

We plan to conduct a series of experiments on search tasks categorized by data scales. These experiments are in the process of development and are expected to reveal the pattern difference on presentations between scale types in information search. Most importantly, they will examine the analyses based on the *UFuRT* design and evaluation framework.

Table 2. Task analysis for three search tasks in three types of interfaces

Steps **Internal/External=ratio**	**Text** **Holistic**	 **Separate**	**Graph** **Holistic**	 **Separate**	**Mix** **Holistic**	 **Separate**
Task I Are there any abnormal levels of cholesterol in the patient record?	26 14/12=1.17	27 14/13=1.07	3-15 ½=0.50	7-21 ¾=.75	4-26 2/2=1.00	6-28 2/4=0.50
Task II In which month of 2003, the patient's LDL level was abnormal?	28 15/13=1.15	29 15/14=1.07	4-27 2/2=1	8-33 3/5=0.80	7-40 2/5=0.40	8-44 2/6=0.33
Task III Did the patient's triglyceride level drop since the start of his diet treatment?	24 12/12=1.00	26 12/14=0.86	2 0/2=0	3 0/3=0	4 0/4=0	5 0/5=0

6. Discussion

The results of analyses reveal the steps that are needed to be performed in order to reach a goal. For each step in a search task, the information needed to carry out the step can be either internal or external information. According to the analyses, the steps and the information required for each step dependent on an interface for a search task jointly determine the efficiency, task complexity, and the possibility of making errors. This research on internal and external information in information search has both theoretical and practical implications. It reveals how a human information seeker interacts with artifacts in information search tasks when conducted under different distributed conditions. The framework of distributed information search and the search task taxonomy contributes theoretically to the study of information search, distributed cognition, and the disciplines of human-centered computing. The practical contribution is an effective prediction and a better design of search interfaces in consideration of data scale and distributed nature of information.

The preliminary results we present in this paper were obtained from theoretical analyses employing the *UFuRT* design and evaluation framework. In previous studies in other domains, we performed evaluation studies to examine the validity of the theoretical predictions. The results from these evaluations were consistent with the theoretical predications. We are currently performing empirical experiments to evaluate the results from theoretical analysis.

We believe that *UFuRT* is applicable to the design and evaluation of usability of interfaces for information search because it is a useful process that not only provides design guidelines but also generates estimates of representational efficiencies, task complexities, and user behavioral outcomes. In the healthcare practice, physicians, nurses and other users spend a large amount of time reading and searching healthcare data in medical records. The prototypes of distributed information search tools based on human-centered computing and distributed cognition will be applicable in the information search tool design of EHR which will support, facilitate, and enhance the healthcare practice.

7. Acknowledgments

This project was supported by a grant from the Human Centered Computing and Intelligent Systems Program, National Aeronautics and Space Administration, Grant No. NCC 2-1234 and No. NNA04CK49G.

This project was also supported in part by a training fellowship from the W.M. Keck Foundation to the Gulf Coast Consortia through the Keck Center for Computational and Structural Biology.

8. References

1. Gaslikova I. "information seeking in context" and the development of information systems. Information Research 1999;5(1).
2. Dervin B. On studying information seeking methodologically. The implications of connecting metatheory to method. Information Processing and Management 1999;35(6):727-750.
3. Eisenberg MB, Berkowitz RE. Information problem-solving: The big six skills approach to library and information skills instruction. Norwood, NJ: Albex Publishing; 1990.
4. Fidel R, Bruce H. Collaborative information retrieval. In; 2000.
5. Lancaster FW, Warner AJ. Information retrieval today. Arlington, Virginia: Information resources press; 1993.
6. Marchionini G. Interfaces for end-user information seeking. Journal of the American society for information science 1992;42(2):156-163.
7. Zhang J. The interaction of internal and external representations in a problem solving task. In: Proceedings of the Thirteenth Annual Conference of Cognitive Science Society; 1991; Hillsdale: NJ: Erlbaum; 1991.
8. Wright PC, Fields RE, Harrison MD. Analysing human-computer interaction as distributed cognition: The resources model. Human-Computer Interaction 2000;15(1):1-41.
9. Horsky J, Kaufman DR, Patel VL. The cognitive complexity of a provider order entry interface. In: Proceeding of AMIA; 2003; Washington, DC; 2003. p. 294-98.
10. Hollan J, Hutchins E, Kirsh D. Distributed cognition: Toward a new foundation for human-computer interaction research. In: ACM Transactions on Computer-Human Interaction; 2000; 2000.
11. Hutchins E. Cognition in the wild. Cambridge, Massachusetts: Massachusetts Institute of Technology; 1995.
12. Norman DA. Things that make us smart: Defending human attributes in the age of the machine. Massachusetts: Addison-Wesley Perseus; 1993.
13. Jacob R. User interface. In: Ralston A, Hemmendinger D, Reilly E, editors. Encyclopedia of computer science. 4th ed: Macmillan; 2000.
14. Komlodi A. Task management support in information seeking: A case for search histories. Computer in Human Behavior 2004;20:163-184.
15. Grudin J. The computer reaches out: The historical continuity of user interface design. In: Proceedings of the CHI'90 conference on human factors in computer systems; 1990; New York: ACM; 1990. p. 261-266.
16. Hayden K. Information seeking models. In; 2002.
17. Zeng Q, Cimino JJ. Providing multiple views to meet physician information needs. In: Proceedings of the 33rd Hawaii International Conference on System Sciences; 2000; Hawaii; 2000.
18. Bystrom K, Jarvelin K. Task complexity affects information seeking and use. Information Processing and Management 1995(2):191-213.
19. Schneiderman B. Designing the user interface. 3rd ed: Addision Wesley Longman.; 1998.
20. Niedzwiedzka B. A proposed general model of information behaviour. Information research 2003;9(1).
21. Allen T. Managing the flow of the technology. Cambridge: the MIT Press; 1977.
22. Dervin B. An overview of sense-making research: Concepts, methods, and results to date. In: annual meeting of the International Communication Association; 1983; Dallas, TX.; 1983.
23. Ellis D. A behavioural model for information retrieval system design. Journal of Information Science 1989;15(4/5):237-247.
24. Ellis D, Cox D, Hall K. A comparison of the information seeking patterns of researchers in the physical and social sciences. Journal of Documentation 1993;49(4):356-369.
25. Kuhlthau CC. Inside the search process: Information seeking from the user's perspective. Journal of the American Society for Information Science 1991;42(5):361-371.
26. Saracevic T. Modeling interaction in information retrieval (ir): A review and proposal. Proceedings of the American Society for Information Science 1996;33:3-9.
27. Zhang J, Patel VL, Johnson KA, Malin J, Smith JW. Designing human-centered distributed information systems. IEEE intelligent systems 2002;17(5):42-47.
28. Kieras D. Goms models: An approach to rapid usability evaluation. In; 2004.